Proceedings of Machine Learning Research vol xxx:1–23, 2021

Towards Robust Data-Driven Control Synthesis for Nonlinear Systems with Actuation Uncertainty

Andrew J. Taylor[1] AJTAYLOR@CALTECH.EDU
Victor D. Dorobantu[1] VDOROBAN@CALTECH.EDU
Caltech, Pasadena, CA, USA

Sarah Dean[1] DEAN_SARAH@BERKELEY.EDU
Benjamin Recht BRECHT@BERKELEY.EDU
University of California, Berkeley, Berkeley, CA, USA

Yisong Yue YYUE@CALTECH.EDU
Aaron D. Ames AMES@CALTECH.EDU
Caltech, Pasadena, CA, USA

Abstract

Modern nonlinear control theory seeks to endow systems with properties such as stability and safety, and has been deployed successfully across various domains. Despite this success, model uncertainty remains a significant challenge in ensuring that model-based controllers transfer to real world systems. This paper develops a data-driven approach to robust control synthesis in the presence of model uncertainty using Control Certificate Functions (CCFs), resulting in a convex optimization based controller for achieving properties like stability and safety. An important benefit of our framework is nuanced data-dependent guarantees, which in principle can yield sample-efficient data collection approaches that need not fully determine the input-to-state relationship. This work serves as a starting point for addressing important questions at the intersection of nonlinear control theory and non-parametric learning, both theoretical and in application. We validate the proposed method in simulation with an inverted pendulum in multiple experimental configurations.

Keywords: robust control, data-driven, Lyapunov, barriers

1. Introduction

Ensuring properties such as stability and safety is of significant importance in many modern control applications, including autonomous driving, industrial robotics, and aerospace vehicles. In practice, the models used to design these controllers are imperfect, with model uncertainty arising due to unmodeled dynamics and parametric errors. In the presence of such uncertainty, controllers may fail to render systems stable or safe. As real world control systems become increasingly complex, the potential for detrimental modeling errors increases, and thus it is critical to study control synthesis in the presence of uncertainty.

In this work, we consider a data-driven approach for capturing model uncertainty in the control synthesis process using *certificate functions* (Boffi et al., 2020), which provide a relationship between system dynamics, controllers, and desired properties such as stability and safety. Data-driven

. [1] These authors contributed equally to this work.

and machine learning based approaches have shown great promise for controlling systems with an uncertain model or with no model at all (Kober et al., 2013; Shi et al., 2019; Cheng et al., 2019; Lee et al., 2020). Furthermore, the integration of techniques from nonlinear control theory for achieving stability and safety with data-driven methods has become increasingly popular (Aswani et al., 2013; Beckers et al., 2019; Berkenkamp and Schoellig, 2015; Gillula and Tomlin, 2012; Qu et al., 2020). Many of these approaches (Khansari-Zadeh and Billard, 2014; Lederer et al., 2020a; Choi et al., 2020; Cohen and Belta, 2020; Castañeda et al., 2020) rely on the use of certificate functions to provide control theoretic guarantees.

Uncertainty in the effect of actuation remains a major challenge in achieving control-theoretic guarantees with data-driven methods. Many existing approaches assume certainty in how actuation enters the dynamics (Fisac et al., 2018; Umlauft et al., 2018; Zheng et al., 2020), or use structured controllers requiring strong characterizations of this uncertainty (Beckers et al., 2019; Lederer et al., 2020b). Other methods require high coverage datasets with (nearly) complete characterizations of the input-to-state relationship (Berkenkamp et al., 2016, 2017); in practice, collecting such data can be prohibitively costly or damage the system. There remains a need for data-driven approaches for characterizing actuation uncertainty without requiring data that is difficult or infeasible to collect.

Our contribution: The contribution of this work is a novel approach for robust data-driven control synthesis for control-affine systems in the presence of model uncertainty, including actuation uncertainty. We develop this approach through the guarantee-based context of *control certificate functions*, which generalize popular tools from nonlinear control for achieving stability and safety such as Control Lyapunov Functions (Artstein, 1983), Control Barrier Functions (Ames et al., 2014a), and Control Barrier-Lyapunov Functions (Prajna and Jadbabaie, 2004). Data is incorporated into a convex-optimization based control synthesis problem as affine inequality constraints which restrict possible model uncertainties, enabling the choice of robust control inputs over convex uncertainty sets. By taking this approach, rather than requiring a full characterization of how input enters the system, we utilize the affine structure of certificate function dynamics to choose inputs such that actuation uncertainty has little impact on the evolution of the certificate function and does not prevent guarantees of stability and safety. In summary, our results show that good performance can be achieved when training data provides sufficiently dense coverage of the state space with targeted excitation in input directions. We believe that the approach proposed in this paper provides a unique perspective for unifying nonlinear control and non-parametric machine learning that is well positioned to pursue answers to both theoretical and application oriented questions at this intersection.

2. Background

This section provides a review of certificate-based nonlinear control synthesis and an overview of how model uncertainty impacts these synthesis methods.

2.1. Control Certificate Functions

Consider a nonlinear control affine system given by:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}, \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, and $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are locally Lipschitz continuous on $\mathbb{R}^n$. Further assume that the origin is an equilibrium point of the uncontrolled system ($\mathbf{f}(\mathbf{0}) = \mathbf{0}$).

Control-affine systems naturally capture many real-world systems such as robotic (Murray et al., 1994) and automotive systems (Ioannou and Chien, 1993). For simplicity of exposition, in this work we assume that $\mathbf{u}$ may be chosen unbounded. Given a locally Lipschitz continuous state-feedback controller $\mathbf{k} : \mathbb{R}^n \to \mathbb{R}^m$, the closed-loop system dynamics are:

$$\dot{\mathbf{x}} = \mathbf{f}_{\text{cl}}(\mathbf{x}) \triangleq \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{k}(\mathbf{x}). \tag{2}$$

The assumption on local Lipschitz continuity of $\mathbf{f}$, $\mathbf{g}$, and $\mathbf{k}$ implies that $\mathbf{f}_{\text{cl}}$ is locally Lipschitz continuous. Thus for any initial condition $\mathbf{x}_0 \triangleq \mathbf{x}(0) \in \mathbb{R}^n$ there exists a maximum time interval $I(\mathbf{x}_0) = [0, t_{\max})$ such that $\mathbf{x}(t)$ is the unique solution to (2) on $I(\mathbf{x}_0)$ (Perko, 2013). In the case that $\mathbf{f}_{\text{cl}}$ is forward complete, $t_{\max} = \infty$.

The qualitative behavior (such as stability or safety) of the the closed-loop system (2) can be described via the notion of a continuously differentiable *certificate function* $\varphi : \mathbb{R}^n \to \mathbb{R}$. Given a *comparison function* $\alpha : \mathbb{R} \to \mathbb{R}$ (specific to the qualitative behavior of interest), certification is specified as an inequality on the derivative of the certificate function along solutions to the closed-loop system:

$$\dot{\varphi}(\mathbf{x}) = \nabla\varphi(\mathbf{x})^\top \mathbf{f}_{cl}(\mathbf{x}) \leq -\alpha(\varphi(\mathbf{x}))\,. \tag{3}$$

Certifying that this property of the closed-loop system (2) is met may be impossible if the controller $\mathbf{k}$ was not chosen to enforce the desired property. This motivates the following definition of a *Control Certificate Function*:

Definition 1 (Control Certificate Function (CCF)) A continuously differentiable function $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a *Control Certificate Function (CCF)* for (1) with comparison function $\alpha : \mathbb{R} \to \mathbb{R}$ if for all $\mathbf{x} \in \mathbb{R}^n$:

$$\inf_{\mathbf{u}\in\mathbb{R}^m} \underbrace{\nabla\varphi(\mathbf{x})^\top \mathbf{f}(\mathbf{x})}_{L_{\mathbf{f}}\varphi(\mathbf{x})} + \underbrace{\nabla\varphi(\mathbf{x})^\top \mathbf{g}(\mathbf{x})}_{L_{\mathbf{g}}\varphi(\mathbf{x})}\mathbf{u} \leq -\alpha(\varphi(\mathbf{x})). \tag{4}$$

where $L_{\mathbf{f}}\varphi : \mathbb{R}^n \to \mathbb{R}$ and $L_{\mathbf{g}}\varphi : \mathbb{R}^n \to \mathbb{R}^m$. The control-affine nature of the system dynamics are preserved by the CCF, such that the evolution of the certificate function is impacted only by the component of the input in the direction of $L_{\mathbf{g}}\varphi(\mathbf{x})$. Given a CCF φ for (1) and a corresponding comparison function α, we define the point-wise set of all control values that satisfy the inequality in (4):

$$K_{\text{ccf}}(\mathbf{x}) \triangleq \left\{ \mathbf{u} \in \mathbb{R}^m \;\middle|\; \nabla\varphi(\mathbf{x})^\top (\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}) \leq -\alpha(\varphi(\mathbf{x})) \right\}. \tag{5}$$

Any nominal locally Lipschitz continuous controller $\mathbf{k}_d : \mathbb{R}^n \to \mathbb{R}^m$ can be modified to take values in the set $K_{\text{ccf}}(\mathbf{x})$ via the certificate-critical CCF-QP:

$$\begin{aligned} \mathbf{k}(\mathbf{x}) = \; & \underset{\mathbf{u}\in\mathbb{R}^m}{\operatorname{argmin}} \quad \frac{1}{2}\|\mathbf{u} - \mathbf{k}_d(\mathbf{x})\|_2^2 \\ & \text{s.t.} \qquad \nabla\varphi(\mathbf{x})^\top (\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}) \leq -\alpha(\varphi(\mathbf{x})). \end{aligned} \tag{CCF-QP}$$

Before providing particular examples of certificate functions useful in control synthesis, we review the following definitions. We denote a continuous function $\alpha : [0, a) \to \mathbb{R}_+$, with $a > 0$, as *class* $\mathcal{K}$ ($\alpha \in \mathcal{K}$) if $\alpha(0) = 0$ and α is strictly monotonically increasing. If $a = \infty$ and $\lim_{r\to\infty} \alpha(r) = \infty$, then α is *class* $\mathcal{K}_\infty$ ($\alpha \in \mathcal{K}_\infty$). A continuous function $\alpha : (-b, a) \to \mathbb{R}$, with $a, b > 0$, is *extended class* $\mathcal{K}$ ($\alpha \in \mathcal{K}_e$) if $\alpha(0) = 0$ and α is strictly monotonically increasing.

If $a, b = \infty$, $\lim_{r\to\infty} \alpha(r) = \infty$, and $\lim_{r\to-\infty} \alpha(r) = -\infty$, then α is *extended class* $\mathcal{K}_\infty$ ($\alpha \in \mathcal{K}_{\infty,e}$). Finally, we note that $c \in \mathbb{R}$ is referred to as a *regular value* of a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$ if $h(\mathbf{x}) = c \implies \nabla h(\mathbf{x}) \neq \mathbf{0}_n$.

Example 1 (Stability via Control Lyapunov Functions) In the context of stabilization to the origin, a control certificate function $V : \mathbb{R}^n \to \mathbb{R}$ with a class $\mathcal{K}$ comparison function $\alpha \in \mathcal{K}$ that satisfies:

$$\alpha_1(\|\mathbf{x}\|) \leq V(\mathbf{x}) \leq \alpha_2(\|\mathbf{x}\|), \tag{6}$$

for $\alpha_1, \alpha_2 \in \mathcal{K}$, is a *Control Lyapunov Function (CLF)* (Artstein, 1983; Sontag, 1989), with stabilization to the origin achieved by controllers taking values in the point-wise set K_{ccf} given by (5) (Ames et al., 2017).

Example 2 (Safety via Control Barrier Functions) In the context of safety, defined as forward invariance (Blanchini, 1999) of a set $\mathcal{C}$, a control certificate function $h : \mathbb{R}^n \to \mathbb{R}$ with 0 a regular value and a comparison function $\alpha \in \mathcal{K}_e^\infty$ that satisfies:

$$\mathbf{x} \in \mathcal{C} \implies h(\mathbf{x}) \leq 0, \tag{7}$$

is a *Control Barrier Function (CBF)* (Ames et al., 2014a, 2017), with safety of the set $\mathcal{C}$ achieved by controllers taking values in the point-wise set K_{ccf} given by (5) (Ames et al., 2019). Note that we adopt the opposite sign convention for h so satisfying (4) guarantees safety.

CLFs and CBFs have been successfully deployed in the context of bipedal robotics (Ames et al., 2014b; Nguyen et al., 2016), adaptive cruise control (Ames et al., 2014a), robotic manipulators (Khansari-Zadeh and Billard, 2014), and multi-agent systems (Pickem et al., 2017).

2.2. Model Uncertainty

In practice, uncertainty in the dynamics (1) of the system exists due to parametric error and unmodeled dynamics, such that the functions $\mathbf{f}$ and $\mathbf{g}$ are not precisely known. Control affine systems are a natural place to study actuation uncertainty as uncertainty in $\mathbf{g}$ can be interpreted as uncertainty in a linear gain multiplying the input. In this setting, control synthesis is done with a nominal model that estimates the true dynamics of the system:

$$\dot{\widehat{\mathbf{x}}} = \widehat{\mathbf{f}}(\mathbf{x}) + \widehat{\mathbf{g}}(\mathbf{x})\mathbf{u}, \tag{8}$$

where $\widehat{\mathbf{f}} : \mathbb{R}^n \to \mathbb{R}^n$ and $\widehat{\mathbf{g}} : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are locally Lipschitz continuous. Adding and subtracting this expression to and from (1), we see that the system evolves under the differential equation:

$$\dot{\mathbf{x}} = \widehat{\mathbf{f}}(\mathbf{x}) + \widehat{\mathbf{g}}(\mathbf{x})\mathbf{u} + \underbrace{\mathbf{f}(\mathbf{x}) - \widehat{\mathbf{f}}(\mathbf{x})}_{\widetilde{\mathbf{f}}(\mathbf{x})} + \underbrace{(\mathbf{g}(\mathbf{x}) - \widehat{\mathbf{g}}(\mathbf{x}))}_{\widetilde{\mathbf{g}}(\mathbf{x})}\mathbf{u}, \tag{9}$$

where $\widetilde{\mathbf{f}} : \mathbb{R}^n \to \mathbb{R}^n$ and $\widetilde{\mathbf{g}} : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are the unmodeled dynamics. This uncertainty in the dynamics additionally manifests in the time derivative of a CCF designed for the system:

$$\dot{\varphi}(\mathbf{x},\mathbf{u}) = \overbrace{\underbrace{\nabla\varphi(\mathbf{x})^\top \widehat{\mathbf{f}}(\mathbf{x})}_{L_{\widehat{\mathbf{f}}}\varphi(\mathbf{x})} + \underbrace{\nabla\varphi(\mathbf{x})^\top \widehat{\mathbf{g}}(\mathbf{x})}_{L_{\widehat{\mathbf{g}}}\varphi(\mathbf{x})}\mathbf{u}}^{\widehat{\dot{\varphi}}(\mathbf{x},\mathbf{u})} + \underbrace{\nabla\varphi(\mathbf{x})^\top \widetilde{\mathbf{f}}(\mathbf{x})}_{L_{\widetilde{\mathbf{f}}}\varphi(\mathbf{x})} + \underbrace{\nabla\varphi(\mathbf{x})^\top \widetilde{\mathbf{g}}(\mathbf{x})}_{L_{\widetilde{\mathbf{g}}}\varphi(\mathbf{x})}\mathbf{u}, \tag{10}$$

where $L_{\widehat{\mathbf{f}}}\varphi, L_{\widetilde{\mathbf{f}}}\varphi : \mathbb{R}^n \to \mathbb{R}$, and $L_{\widehat{\mathbf{g}}}\varphi, L_{\widetilde{\mathbf{g}}}\varphi : \mathbb{R}^n \to \mathbb{R}^m$. The presence of uncertainty in the CCF time derivative makes it impossible to verify whether a given control input satisfies the inequality in (4), and can lead to failure to achieve the desired qualitative behavior. This model decomposition reveals an important feature of the actuation uncertainty in that if $L_{\widetilde{\mathbf{g}}}\varphi$ can be restricted such that there exists a direction $\mathbf{s} \in \mathbb{R}^m$ with $L_{\widehat{\mathbf{g}}}\varphi\mathbf{s} > L_{\widetilde{\mathbf{g}}}\varphi\mathbf{s}$, then by the unboundedness of $\mathbf{u}$, the value of the true CCF time derivative can be made to satisfy the CCF condition (4) by choosing inputs along the direction $-\mathbf{s}$. This may be interpreted as choosing inputs along a direction for which the infinite gain margin property inherent in CCF controllers holds (Jankovic, 2018). Two assumptions that we make on the uncertainty in the dynamics are as follows:

Assumption 1 The function $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a valid CCF with comparison function $\alpha : \mathbb{R} \to \mathbb{R}$ for the uncertain dynamic system (9). Mathematically this assumption appears as:

$$\inf_{\mathbf{u}\in\mathbb{R}^m} \nabla\varphi(\mathbf{x})^\top \left(\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}\right) \leq -\alpha(\varphi(\mathbf{x})).$$

Assumption 2 The functions $\widetilde{\mathbf{f}}$ and $\widetilde{\mathbf{g}}$ are locally Lipschitz continuous with known (but not necessarily strict) Lipschitz constants $\mathcal{L}_{\widetilde{\mathbf{f}}}$ and $\mathcal{L}_{\widetilde{\mathbf{g}}}$.

Intuitively the first assumption implies the control problem is well-posed in that the desired qualitative behavior can be achieved on the uncertain system dynamics (9), though the presence of model-uncertainty makes it impossible to verify if the inequality (4) is met for a given controller. The validity of this assumption has been demonstrated in multiple settings with reasonable model uncertainty (Taylor et al., 2019, 2020), and is made implicitly when transferring these control approaches to real systems. The local Lipschitz continuity in the second assumption follows from the assumption of local Lipschitz continuity of the true dynamics (1) and nominal model (8). Knowledge of strict Lipschitz constants is not necessary, but tighter Lipschitz constants are associated with improvements in the performance of data-driven robust control methods. Approaches for learning Lipschitz coefficients from data have also recently been considered (Fazlyab et al., 2019).

3. Data-Driven Robust Control Synthesis

In this section we explore how data can be incorporated directly into an optimization based controller to robustly achieve a desired qualitative behavior specified via a Control Certificate Function.

Consider a dataset consisting of N tuples of states, inputs, and corresponding state time derivatives, $\mathfrak{D} = \{(\mathbf{x}_i, \mathbf{u}_i, \dot{\mathbf{x}}_i)\}_{i=1}^N$, with $\mathbf{x}_i \in \mathbb{R}^n$, $\mathbf{u}_i \in \mathbb{R}^m$, and $\dot{\mathbf{x}}_i \in \mathbb{R}^n$ for $i = 1, \ldots, N$. It may not be possible to directly measure the state time derivatives $\dot{\mathbf{x}}_i$, but they can approximated from sequential state observations $\mathbf{x}_i$ (Taylor et al., 2019, 2020). For simplicity of exposition, we do not consider noise in this formulation. However, the resulting construction may be modified to account for the impact of bounded noise in the data. Considering the uncertain model (9) evaluated at a state and input pair $(\mathbf{x}_i, \mathbf{u}_i)$ in the data set yields:

$$\widetilde{\mathbf{F}}_i \triangleq \dot{\mathbf{x}}_i - (\widehat{\mathbf{f}}(\mathbf{x}_i) + \widehat{\mathbf{g}}(\mathbf{x}_i)\mathbf{u}_i) = \widetilde{\mathbf{f}}(\mathbf{x}_i) + \widetilde{\mathbf{g}}(\mathbf{x}_i)\mathbf{u}_i, \tag{11}$$

where $\widetilde{\mathbf{F}}_i \in \mathbb{R}^n$ can be interpreted as the error between the true state time derivative and the nominal model (8) evaluated at the state and input pair $(\mathbf{x}_i, \mathbf{u}_i)$.

Considering a state $\mathbf{x} \in \mathbb{R}^n$ (that is not necessarily part of the data set $\mathfrak{D}$) and noting that by the second equality in (11) it follows that:

$$\widetilde{\mathbf{f}}(\mathbf{x}) + \widetilde{\mathbf{g}}(\mathbf{x})\mathbf{u}_i - \widetilde{\mathbf{F}}_i = \widetilde{\mathbf{f}}(\mathbf{x}) - \widetilde{\mathbf{f}}(\mathbf{x}_i) + (\widetilde{\mathbf{g}}(\mathbf{x}) - \widetilde{\mathbf{g}}(\mathbf{x}_i))\,\mathbf{u}_i. \tag{12}$$

This expression provides a relationship between the possible values of the unmodeled dynamics $\widetilde{\mathbf{f}}$ and $\widetilde{\mathbf{g}}$ at the state $\mathbf{x}$ and the values of the unmodeled dynamics at the data point $\mathbf{x}_i$. Using the local Lipschitz continuity of the unmodeled dynamics yields the following bound:

$$\begin{aligned}\left\|\widetilde{\mathbf{f}}(\mathbf{x}) + \widetilde{\mathbf{g}}(\mathbf{x})\mathbf{u}_i - \widetilde{\mathbf{F}}_i\right\|_2 &= \left\|\widetilde{\mathbf{f}}(\mathbf{x}) - \widetilde{\mathbf{f}}(\mathbf{x}_i) + (\widetilde{\mathbf{g}}(\mathbf{x}) - \widetilde{\mathbf{g}}(\mathbf{x}_i))\,\mathbf{u}_i\right\|_2, \\ &\leq \left(\mathcal{L}_{\widetilde{\mathbf{f}}} + \mathcal{L}_{\widetilde{\mathbf{g}}}\|\mathbf{u}_i\|_2\right)\|\mathbf{x} - \mathbf{x}_i\|_2 \triangleq \epsilon_i(\mathbf{x}).\end{aligned} \tag{13}$$

We see that the bound grows with the magnitude of the Lipschitz constants $\mathcal{L}_{\widetilde{\mathbf{f}}}$ and $\mathcal{L}_{\widetilde{\mathbf{g}}}$ and distance of the state $\mathbf{x}$ from the data point $\mathbf{x}_i$. The values of $\mathcal{L}_{\widetilde{\mathbf{f}}}$ and $\mathcal{L}_{\widetilde{\mathbf{g}}}$ are not explicitly data dependent, and thus the bound can be improved for a given data set by reducing the possible model uncertainty through improved modeling. Given this construction we may define the point-wise uncertainty set:

$$\mathcal{U}_i(\mathbf{x}) \triangleq \left\{(\mathbf{A}, \mathbf{b}) \in \mathbb{R}^{n\times m} \times \mathbb{R}^n \;\middle|\; \left\|\mathbf{b} + \mathbf{A}\mathbf{u}_i - \widetilde{\mathbf{F}}_i\right\|_2 \leq \epsilon_i(\mathbf{x})\right\} \subset \mathbb{R}^{n\times m} \times \mathbb{R}^n, \tag{14}$$

noting that $(\widetilde{\mathbf{g}}(\mathbf{x}), \widetilde{\mathbf{f}}(\mathbf{x})) \in \mathcal{U}_i(\mathbf{x})$ and $\mathcal{U}_i(\mathbf{x})$ is convex. Considering this construction over the entire data set $\mathfrak{D}$ we define:

$$\mathcal{U}(\mathbf{x}) \triangleq \bigcap_{i=1}^{N} \mathcal{U}_i(\mathbf{x}) \subset \mathbb{R}^{n\times m} \times \mathbb{R}^n, \tag{15}$$

noting that $(\widetilde{\mathbf{g}}(\mathbf{x}), \widetilde{\mathbf{f}}(\mathbf{x})) \in \mathcal{U}(\mathbf{x})$ and $\mathcal{U}(\mathbf{x})$ is convex. Therefore, $\mathcal{U}(\mathbf{x})$ consists of all possible model errors that are consistent with the observed data. This allows us to pose the following data robust control problem:

Definition 2 (Data Robust Control Certificate Function Optimization Problem)

$$\begin{aligned}\mathbf{k}_{\text{rob}}(\mathbf{x}) = \underset{\mathbf{u}\in\mathbb{R}^m}{\operatorname{argmin}}\ & \frac{1}{2}\|\mathbf{u} - \mathbf{k}_d(\mathbf{x})\|_2^2 \\ \text{s.t. } & \dot{\widehat{\varphi}}(\mathbf{x}, \mathbf{u}) + \nabla\varphi(\mathbf{x})^\top(\mathbf{b} + \mathbf{A}\mathbf{u}) \leq -\alpha(\varphi(\mathbf{x})) \ \text{ for all } (\mathbf{A}, \mathbf{b}) \in \mathcal{U}(\mathbf{x}).\end{aligned} \tag{DR-CCF-OP}$$

By construction we have that $(\widetilde{\mathbf{g}}(\mathbf{x}), \widetilde{\mathbf{f}}(\mathbf{x})) \in \mathcal{U}(\mathbf{x})$, implying that $\mathbf{k}_{\text{rob}}(\mathbf{x}) \in K_{\text{ccf}}(\mathbf{x})$ when the problem is feasible. Thus the closed-loop system (2) under $\mathbf{k}_{\text{rob}}$ satisfies inequality (3). We next present one of our main results, using robust optimization (Ben-Tal et al., 2009) to yield a convex problem for synthesizing such a robust controller, proven in the Appendix A

Theorem 1 (Robust Control Synthesis) Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a control certificate with comparison function $\alpha : \mathbb{R} \to \mathbb{R}$. The robust controller (DR-CCF-OP) can be equivalently expressed as:

$$\begin{aligned}\mathbf{k}_{\text{rob}}(\mathbf{x}) = \underset{\substack{\mathbf{u}\in\mathbb{R}^m \\ \boldsymbol{\lambda}_i\in\mathbb{R}^n}}{\operatorname{argmin}}\ & \frac{1}{2}\|\mathbf{u} - \mathbf{k}_d(\mathbf{x})\|_2^2 \\ \text{s.t. } & \dot{\widehat{\varphi}}(\mathbf{x}, \mathbf{u}) - \sum_{i=1}^{N}\left(\boldsymbol{\lambda}_i^\top \widetilde{\mathbf{F}}_i - \|\boldsymbol{\lambda}_i\|_2\epsilon_i(\mathbf{x})\right) \leq -\alpha(\varphi(\mathbf{x})), \\ & \sum_{i=1}^{N}\boldsymbol{\lambda}_i\mathbf{u}_i^\top = -\nabla\varphi(\mathbf{x})\mathbf{u}^\top, \quad \sum_{i=1}^{N}\boldsymbol{\lambda}_i = -\nabla\varphi(\mathbf{x}).\end{aligned} \tag{DR-CCF-SOCP}$$

4. Analysis

In this section we provide an analysis of features of the controller proposed in Section 3.

Convexity & Dimensionality: The inequality constraint in (DR-CCF-SOCP) is convex in each $\boldsymbol{\lambda}_i$ since $\epsilon_i(\mathbf{x}) \geq 0$. As the remainder of the constraints are affine in $\mathbf{u}$ and each $\boldsymbol{\lambda}_i$ and the cost function is convex, this optimization problem is convex. Although this means that a global optimizer can be found efficiently, we note that the problem is high-dimensional, as each data point used in the construction of the optimization problem adds n decision variables via $\boldsymbol{\lambda}_i$. However, if the value of $\epsilon_i(\mathbf{x})$ is large for a given data point (i.e. the data point $\mathbf{x}_i$ is far from the evaluation point $\mathbf{x}$), the magnitude of $\boldsymbol{\lambda}_i$ will be vanishingly small so that the inequality constraint is met. This suggests that the dimensionality of this controller can be reduced by using subsets of the data that are spatially close to the evaluation point. It will therefore be important to develop a mathematical understanding of which data point are most relevant to the controller in regions of the state space, and build data structures to segment data such that it can be efficiently accessed by the closed-loop controller, such as in work by Lederer et al. (2020c).

Feasibility: The feasibility of the the controller (DR-CCF-SOCP) at a given state $\mathbf{x}$ is determined by the structure of the uncertainty set $\mathcal{U}(\mathbf{x})$ defined in (15). The following lemma provides a condition on the diversity of inputs in the data set that implies $\mathcal{U}(\mathbf{x})$ is bounded for all $\mathbf{x} \in \mathbb{R}^n$:

Lemma 1 (Bounded Uncertainty Sets) Consider the data set $\mathfrak{D}$ with N data points satisfying $N \geq m+1$. If there exists a set of data points $\{(\mathbf{x}_i, \mathbf{u}_i, \dot{\mathbf{x}}_i)\}_{i=1}^{m+1} \subseteq \mathfrak{D}$ such that the set of vectors $\mathcal{S} \triangleq \left\{ \begin{bmatrix} \mathbf{u}_i^\top & 1 \end{bmatrix}^\top \right\}_{i=1}^{m+1}$ are linearly independent, then the uncertainty set $\mathcal{U}(\mathbf{x})$ is bounded (and thus compact) for any $\mathbf{x} \in \mathbb{R}^n$.

Again, this result shows that input diversity is sufficient to assert boundedness of the uncertainty set as a uniform property over the entire state space. As seen in the proof, the bound on this set may be very large if the values of $\epsilon_i(\mathbf{x})$ are large (as is the case when $\mathbf{x}$ is far away from the data points $\mathbf{x}_i$ associated with the inputs used to construct the bound). Alternatively, the uncertainty set will be small for a point $\mathbf{x}$ if there is local input diversity within the training dataset. To understand how the size of $\mathcal{U}(\mathbf{x})$ impacts feasibility of the optimization problem, we define the following sets:

$$\widetilde{\mathcal{U}}_\varphi(\mathbf{x}) \triangleq \left\{ (\mathbf{a}, b) \in \mathbb{R}^m \times \mathbb{R} \mid \exists\, (\mathbf{A}, \mathbf{b}) \in \mathcal{U}(\mathbf{x}) \text{ s.t. } \mathbf{a} = (\nabla\varphi(\mathbf{x})^\top \mathbf{A})^\top,\ b = \nabla\varphi(\mathbf{x})^\top \mathbf{b} \right\}, \quad (16)$$

$$\mathcal{U}_\varphi(\mathbf{x}) \triangleq \left\{ \left(L_{\widehat{\mathbf{g}}}\varphi(\mathbf{x})^\top, L_{\widehat{\mathbf{f}}}\varphi(\mathbf{x}) \right) \right\} \oplus \widetilde{\mathcal{U}}_\varphi(\mathbf{x}), \quad (17)$$

where $\oplus$ denotes a Minkowski sum. The set $\widetilde{\mathcal{U}}_\varphi(\mathbf{x})$ can be interpreted as the projection of the dynamics uncertainty set $\mathcal{U}(\mathbf{x})$ along the gradient of the control certificate function φ, creating an $m+1$ dimensional set which represents the possible uncertainties in the *Lie derivatives* of φ. The set $\mathcal{U}_\varphi(\mathbf{x})$ is the recentering of this set around the estimate of the Lie derivatives $(L_{\widehat{\mathbf{g}}}\varphi(\mathbf{x})^\top, L_{\widehat{\mathbf{f}}}\varphi(\mathbf{x})) \in \mathbb{R}^{m+1}$, such that it captures the possible true Lie derivatives of φ. The continuous differentiability of φ implies that if $\mathcal{U}(\mathbf{x})$ is bounded, then $\widetilde{\mathcal{U}}_\varphi(\mathbf{x})$ and $\mathcal{U}_\varphi(\mathbf{x})$ are bounded. Furthermore, as multiplication by $\nabla\varphi(\mathbf{x})^\top$ is a linear transformation, $\widetilde{\mathcal{U}}_\varphi(\mathbf{x})$ and $\mathcal{U}_\varphi(\mathbf{x})$ are convex. We now state a necessary and sufficient condition for feasibility of the optimization problem:

Theorem 2 (Feasibility of Data-Driven Robust Controller) For a state $\mathbf{x} \in \mathbb{R}^n$, let the sets $\mathcal{U}(\mathbf{x})$, $\widetilde{\mathcal{U}}_\varphi(\mathbf{x})$, and $\mathcal{U}_\varphi(\mathbf{x})$ be defined as in (15), (16), and (17), respectively. Assume that $\mathcal{U}(\mathbf{x})$ is bounded, and define the ray $\mathcal{R} \subset \mathbb{R}^{m+1}$ as $\mathcal{R} = \{\mathbf{0}_m\} \times [-\alpha(\varphi(\mathbf{x})), \infty)$. The data-driven robust controller (DR-CCF-SOCP) is feasible at $\mathbf{x} \neq \mathbf{0}$ if and only if:

$$\mathcal{U}_\varphi(\mathbf{x}) \cap \mathcal{R} = \emptyset. \tag{18}$$

We note the controller is trivially feasible at $\mathbf{x} = \mathbf{0}$. A proof of Lemma 1 and Theorem 2 are provided in Appendix B. Intuitively, the ray $\mathcal{R}$ represents Lie derivative pairs that do not satisfy the certificate function condition with no actuation and cannot be modified through actuation. By Assumption 1, the Lie derivatives of the true system are not contained in $\mathcal{R}$, but the possible uncertainties permitted by data need not necessarily reflect this. If one possible uncertainty pair in the set $\mathcal{U}_\varphi(\mathbf{x})$ is contained in $\mathcal{R}$, it is impossible to meet the certificate condition for that uncertainty pair.

5. Simulation

To demonstrate the behavior of the proposed controller, we run simulated experiments in the setting of a inverted pendulum. The true system has modified inverted pendulum dynamics with $\mathbf{x} = [\theta, \dot{\theta}]$, $m = \ell = 0.7$, and a state-dependent input gain given by $1 - 0.75\exp(-\theta^2)$. The input attenuation is highest when the pendulum is vertical. We explore data-driven stability and safety with this system for different methods of gathering training data. We compare the robust data-driven controller with an oracle controller, which knows the true model, and a nominal controller, which treats the estimated model as if it were true. In each setting, the estimated model mis-specifies the pendulum mass and length and assumes that the input gain is 1. As a result, the Lipschitz constants of the errors can be bounded by $\mathcal{L}_{\tilde{\mathbf{f}}} = g|\ell - \hat{\ell}|/(\ell\hat{\ell})$ and $\mathcal{L}_{\tilde{\mathbf{g}}} = 0.75\exp(-0.25)/(m\ell^2)$. The certificate functions of interest are quadratic functions with parameter $P = [[1.73205081, 1], [1, 1.73205081]]$. In all settings, the comparison function is linear with $\alpha(r) = \lambda_{\min}(P)r$. All simulations are integrated with Runge-Kutta 4(5) at 100 Hz.

Stabilizing with dense data. We first consider the task of stabilization with $\varphi(\mathbf{x}) = \mathbf{x}^\top P\mathbf{x}$ and $\mathbf{k}_d(\mathbf{x})$ given by a feedback linearizing controller designed with the estimated system. In this setting, the estimated system underestimates the mass and length by 10%. Training data is collected as a grid over θ and $\dot{\theta}$ with grid sizes $\epsilon_\theta = 6.25 \times 10^{-3}$ and $\epsilon_{\dot{\theta}} = 1.25 \times 10^{-2}$. At each point on this grid

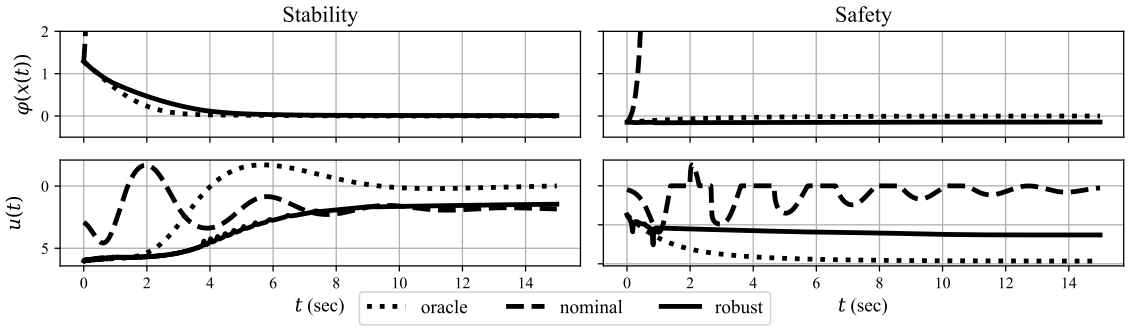

Figure 1. The value of the certificate function $\varphi(\mathbf{x})$ and input $\mathbf{u}$ over time for the dense data settings.

in state space, we sample two control actions $u = \pm 1$. The derivatives $\dot{\mathbf{x}}$ are computed exactly. We include a slack variable in the robust optimization problem for numerical stability. The left panel of Figure 1 displays a system rollout starting at $\mathbf{x}(0) = [0.8, 0.1]$. The nominal controller fails to stabilize the system, as evident by the non-decreasing certificate function. However, the robust controller does so successfully.

Safety with dense data. We next consider the task of ensuring safety via the invariance of an ellipsoidal set with $\varphi(\mathbf{x}) = \mathbf{x}^\top P \mathbf{x} - 0.2$ and $\mathbf{k}_d(\mathbf{x}) = 0$. We suppose that the pendulum length and mass are underestimated to $\widehat{\ell} = \widehat{m} = 0.5$, or about 28%. As above, training data is collected in a grid with positive and negative inputs. The grid is not as dense, with $\epsilon_\theta = 1.25 \times 10^{-2}$ and $\epsilon_{\dot{\theta}} = 5 \times 10^{-2}$. The right panel in Figure 1 displays a system rollout starting within the safe set at $\mathbf{x}(0) = [0.1, 0.1]$. The nominal controller fails to keep the system safe, as evident by the positive certificate function. The robust controller successfully maintains safety, though it converges to a different point than the oracle controller.

Safety with trajectory data. In this setting, we consider the same safety task as above with the pendulum mass and length underestimated by 10%. Rather than using a grid, we collect training data from closed-loop operation. We construct a training set using a time-varying proportional control law, $u(\theta, t) = -(25 + \sin(2\pi t/100))\theta$. This controller results in trajectories which are oscillatory and unstable. Training data consists of four trajectories with initial conditions at $[\theta_0, 0]$ for $\theta_0 \in \{0.07, 0.08, 0.09, 0.1\}$, and the derivatives $\dot{\mathbf{x}}$ are computed using numerical differentiation along the recorded trajectory. The resulting data is displayed the right panel of Figure 2. The left panel of Figure 2 displays the timeseries for a system rollout starting at $\mathbf{x}(0) = [0.1, 0.1]$, while the right panel displays the phase plot. As before, the nominal controller fails to keep the system safe. The robust data-driven controller successfully ensures safety, though it again converges to a different point than the oracle controller. This setting demonstrates that data entirely from unsafe and unstable trajectories provides sufficient information to synthesize a safe controller.

Safety with episodic data. Finally, we consider the above safety task performed over multiple episodes, collecting and aggregating data after each. For this setting, we remove the exponential attenuation of the input gain for the true system, so $\mathcal{L}_{\widetilde{\mathbf{g}}} = 0$. The training data is initialized using the

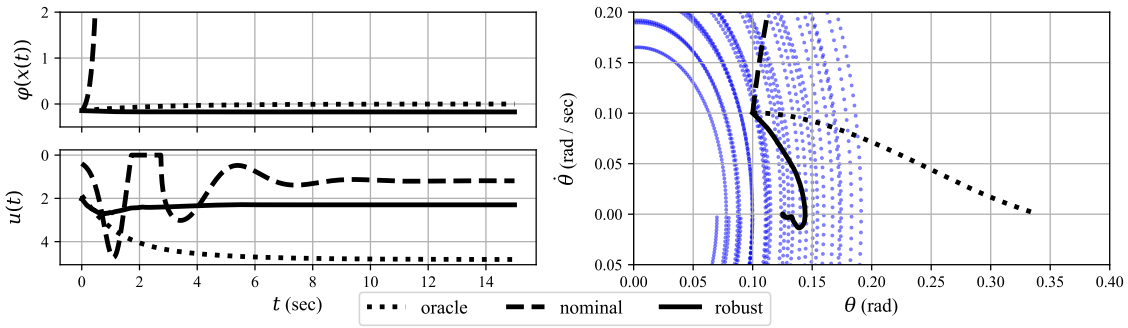

Figure 2. (Left) The value of the certificate function $\varphi(\mathbf{x})$ and input $\mathbf{u}$ over time for the trajectory data setting. (Right) The corresponding trajectories in state space and the training data (blue).

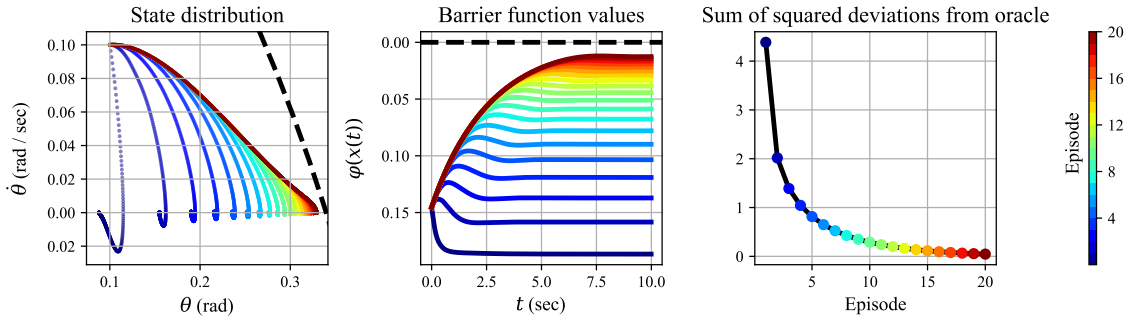

Figure 3. (Left) The state distribution observed during each episode and the boundary of the safe set (black dashed). (Center) The value of the certificate function $\varphi(\mathbf{x})$ over time for each episode and the maximum permissible value (black dashed). (Right) The total additional input compensating for remaining uncertainty during each episode.

proportional control law $u(\theta) = -25\theta$, starting from the initial condition $[0.1, 0]$. The derivatives $\dot{\mathbf{x}}$ are computed exactly. We collect data from a system rollout starting at $\mathbf{x}(0) = [0.1, 0.1]$ using the robust controller with the initial training data. The observed trajectory is subsampled by a factor of 10 and added to the training data. We repeat rollouts and data aggregation for a total of 20 episodes. Aggregation of state data is shown in the left panel of Figure 3, and the value of the certificate function during each episode is shown in the center panel of Figure 3. In early episodes, the robust controller does not have enough data to safely allow the system to approach the boundary of the safe set. As the episodes progress and data is collected in regions of the state space closer to the boundary, the reduced uncertainty in these regions allows the system to approach the boundary. We compare the actions of the robust controller to the actions that the oracle controller would have taken on the same trajectories, measuring the total additional input compensating for uncertainty over an episode; this measurement is shown as a function of number of episodes in the right panel of Figure 3. As more data is collected, the performance of the robust controller more closely matches that of the oracle controller. Importantly, this occurs while providing the robust controller only with additional data inside the safe set collected with a safe controller; we do not have to collect data outside the safe set or with unsafe excitations to exhibit these changes.

6. Future Work

In this paper, we propose a novel approach for robust data-driven control synthesis in the presence of model uncertainty and show that good performance can be achieved when training data provides sufficiently dense coverage of the state space with targeted excitation in input directions. We believe that this approach is well positioned to make progress on both theoretical and application-oriented problems at the intersection nonlinear control and non-parametric machine learning. There are several open directions for future work. To enable efficient computation for closed-loop control, it is important to investigate strategies for data organization and segmentation, which may arise from a careful analysis of Lagrange multiplier support. It is also important to understand how data collection can be targeted to tasks, and make connections to imitation learning. Finally, it is important to understand properties of the resulting controller, such as continuity, and develop practical implementation considerations for use in real systems.

References

A. Ames, J. Grizzle, and P. Tabuada. Control barrier function based quadratic programs with application to adaptive cruise control. In *Conference on Decision & Control (CDC)*, pages 6271–6278. IEEE, 2014a.

A. D. Ames, K. Galloway, K. Sreenath, and J. W. Grizzle. Rapidly exponentially stabilizing control lyapunov functions and hybrid zero dynamics. *Transactions on Automatic Control*, 59(4):876–891, 2014b.

A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada. Control barrier function based quadratic programs for safety critical systems. *Transactions on Automatic Control*, 62(8):3861–3876, 2017.

A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada. Control barrier functions: Theory and applications. In *European Control Conference (ECC)*, pages 3420–3431. IEEE, 2019.

Z. Artstein. Stabilization with relaxed controls. *Nonlinear Analysis: Theory, Methods & Applications*, 7(11):1163–1173, 1983.

A. Aswani, H. Gonzalez, S. S. Sastry, and C. Tomlin. Provably safe and robust learning-based model predictive control. *Automatica*, 49(5):1216–1226, 2013.

T. Beckers, D. Kulić, and S. Hirche. Stable gaussian process based tracking control of euler–lagrange systems. *Automatica*, 103:390–397, 2019.

A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*, volume 28. Princeton University Press, 2009.

F. Berkenkamp and A. P. Schoellig. Safe and robust learning control with gaussian processes. In *European Control Conference (ECC)*, pages 2496–2501. IEEE, 2015.

F. Berkenkamp, R. Moriconi, A. P. Schoellig, and A. Krause. Safe learning of regions of attraction for uncertain, nonlinear systems with gaussian processes. In *Conference on Decision & Control (CDC)*, pages 4661–4666. IEEE, 2016.

F. Berkenkamp, M. Turchetta, A. Schoellig, and A. Krause. Safe model-based reinforcement learning with stability guarantees. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 908–918, 2017.

F. Blanchini. Set invariance in control. *Automatica*, 35(11):1747–1767, 1999.

N. M. Boffi, S. Tu, N. Matni, J.-J. E. Slotine, and V. Sindhwani. Learning stability certificates from data. *arXiv preprint arXiv:2008.05952*, 2020.

S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

F. Castañeda, J. J. Choi, B. Zhang, C. J. Tomlin, and K. Sreenath. Gaussian process-based min-norm stabilizing controller for control-affine systems with uncertain input effects. *arXiv preprint arXiv:2011.07183*, 2020.

R. Cheng, G. Orosz, R. M. Murray, and J. W. Burdick. End-to-end safe reinforcement learning through barrier functions for safety-critical continuous control tasks. In *Conference on Artificial Intelligence*, volume 33, pages 3387–3395. AAAI, 2019.

J. Choi, F. Castañeda, C. Tomlin, and K. Sreenath. Reinforcement Learning for Safety-Critical Control under Model Uncertainty, using Control Lyapunov Functions and Control Barrier Functions. In *Robotics: Science and Systems*, Corvalis, Oregon, USA, July 2020. doi: 10.15607/RSS.2020.XVI.088.

M. Cohen and C. Belta. Approximate optimal control for safety-critical systems with control barrier functions. *arXiv preprint arXiv:2008.04122*, 2020.

M. Fazlyab, A. Robey, H. Hassani, M. Morari, and G. Pappas. Efficient and accurate estimation of lipschitz constants for deep neural networks. In *Advances in Neural Information Processing Systems (NeurIPS*, pages 11427–11438, 2019.

J. F. Fisac, A. K. Akametalu, M. N. Zeilinger, S. Kaynama, J. Gillula, and C. J. Tomlin. A general safety framework for learning-based control in uncertain robotic systems. *Transactions on Automatic Control*, 2018.

J. H. Gillula and C. J. Tomlin. Guaranteed safe online learning via reachability: tracking a ground target using a quadrotor. In *International Conference on Robotics and Automation (ICRA)*, pages 2723–2730. IEEE, 2012.

P. A. Ioannou and C.-C. Chien. Autonomous intelligent cruise control. *IEEE Transactions on Vehicular technology*, 42(4):657–672, 1993.

M. Jankovic. Robust control barrier functions for constrained stabilization of nonlinear systems. *Automatica*, 96:359–367, 2018.

S. M Khansari-Zadeh and A. Billard. Learning control lyapunov function to ensure stability of dynamical system-based robot reaching motions. *Robotics and Autonomous Systems*, 62(6), 2014.

J. Kober, J. A. Bagnell, and J. Peters. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, 32(11):1238–1274, 2013.

A. Lederer, A. Capone, and S. Hirche. Parameter optimization for learning-based control of control-affine systems. *Proceedings of Machine Learning Research (PMLR)*, 120:465–475, 2020a.

A. Lederer, A. Capone, J. Umlauft, and S. Hirche. How training data impacts performance in learning-based control. *arXiv preprint arXiv:2005.12062*, 2020b.

A. Lederer, A. J. O. Conejo, K. Maier, W. Xiao, and S. Hirche. Real-time regression with dividing local gaussian processes. *arXiv preprint arXiv:2006.09446*, 2020c.

J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter. Learning quadrupedal locomotion over challenging terrain. *Science robotics*, 5(47), 2020.

R. M. Murray, Z. Li, and S. S. Sastry. *A mathematical introduction to robotic manipulation*. CRC press, 1994.

Q. Nguyen, A. Hereid, J. W. Grizzle, A. D. Ames, and K. Sreenath. 3d dynamic walking on stepping stones with control barrier functions. In *Conference on Decision and Control (CDC)*, pages 827–834. IEEE, 2016.

L. Perko. *Differential equations and dynamical systems*, volume 7. Springer Science & Business Media, 2013.

D. Pickem, P. Glotfelter, L. Wang, M. Mote, A. Ames, E. Feron, and M. Egerstedt. The robotarium: A remotely accessible swarm robotics research testbed. In *International Conference on Robotics and Automation (ICRA)*, pages 1699–1706. IEEE, 2017.

S. Prajna and A. Jadbabaie. Safety verification of hybrid systems using barrier certificates. In *International Workshop on Hybrid Systems: Computation and Control (HSCC)*, pages 477–492. Springer, 2004.

G. Qu, C. Yu, S. Low, and A. Wierman. Combining model-based and model-free methods for nonlinear control: A provably convergent policy gradient approach. *arXiv preprint arXiv:2006.07476*, 2020.

G. Shi, X. Shi, M. O'Connell, R. Yu, K. Azizzadenesheli, A. Anandkumar, Y. Yue, and S. Chung. Neural lander: Stable drone landing control using learned dynamics. In *International Conference on Robotics and Automation (ICRA)*, pages 9784–9790. IEEE, 2019.

E. D. Sontag. A 'universal' construction of artstein's theorem on nonlinear stabilization. *Systems & Control Letters*, 13(2):117–123, 1989.

A. J. Taylor, V. D. Dorobantu, H. M. Le, Y. Yue, and A. D. Ames. Episodic learning with control lyapunov functions for uncertain robotic systems. In *International Conference on Intelligent Robots and Systems (IROS)*, pages 6878–6884. IEEE, 2019.

A. J. Taylor, A. Singletary, Y. Yue, and A. D. Ames. Learning for safety-critical control with control barrier functions. *Proceedings of Machine Learning Research (PMLR)*, 120:708–717, 2020.

J. Umlauft, L. Pöhler, and S. Hirche. An uncertainty-based control lyapunov approach for control-affine systems modeled by gaussian process. *IEEE Control Systems Letters*, 2(3):483–488, 2018. doi: 10.1109/LCSYS.2018.2841961.

L. Zheng, J. Pan, R. Yang, H. Cheng, and H. Hu. Learning-based safety-stability-driven control for safety-critical systems under model uncertainties. *arXiv preprint arXiv:2008.03421*, 2020.

Appendix A. Equivalence Proof

In this appendix we provide a proof of Theorem 1 demonstrating the equivalence between the robust optimization problem (DR-CCF-OP) and the second-order cone form (DR-CCF-SOCP).

Proof An input $\mathbf{u} \in \mathbb{R}^m$ is feasible if the optimal value of the optimization problem:

$$\begin{aligned} \sup_{\substack{\mathbf{A}\in\mathbb{R}^{n\times m}\\ \mathbf{b}\in\mathbb{R}^n}} \ & \widehat{\dot{\varphi}}(\mathbf{x},\mathbf{u}) + \nabla\varphi(\mathbf{x})^\top(\mathbf{b}+\mathbf{A}\mathbf{u}) \\ \text{s.t.} \ & \|\mathbf{b}+\mathbf{A}\mathbf{u}_i - \widetilde{\mathbf{F}}_i\|_2 \leq \epsilon_i(\mathbf{x}) \text{ for all } i \in 1,\ldots,N, \end{aligned}$$

is less than or equal to $-\alpha(\varphi(\mathbf{x}))$. Each inequality constraint can be rewritten as set membership in a second-order cone. Thus we can rewrite this optimization problem as:

$$\begin{aligned} \sup_{\substack{\mathbf{A}\in\mathbb{R}^{n\times m}\\ \mathbf{b}\in\mathbb{R}^m}} \ & \widehat{\dot{\varphi}}(\mathbf{x},\mathbf{u}) + \nabla\varphi(\mathbf{x})^\top(\mathbf{b}+\mathbf{A}\mathbf{u}) \\ \text{s.t.} \ & \left(\mathbf{b}+\mathbf{A}\mathbf{u}_i - \widetilde{\mathbf{F}}_i, \epsilon_i(\mathbf{x})\right) \in \mathcal{Q}_n \text{ for all } i \in 1,\ldots,N, \end{aligned}$$

where $\mathcal{Q}_n \subset \mathbb{R}^{n+1}$ denotes the Lorentz cone $\mathcal{Q}_n = \{(\mathbf{y},t) \in \mathbb{R}^n \times \mathbb{R} : \|\mathbf{y}\|_2 \leq t\}$. The associated Lagrangian for this optimization problem is given by:

$$\begin{aligned} &\widehat{\dot{\varphi}}(\mathbf{x},\mathbf{u}) + \nabla\varphi(\mathbf{x})^\top(\mathbf{b}+\mathbf{A}\mathbf{u}) + \sum_{i=1}^N \begin{bmatrix}\boldsymbol{\lambda}_i\\ \nu_i\end{bmatrix}^\top \begin{bmatrix}\mathbf{b}+\mathbf{A}\mathbf{u}_i - \widetilde{\mathbf{F}}_i\\ \epsilon_i(\mathbf{x})\end{bmatrix} \\ &= \widehat{\dot{\varphi}}(\mathbf{x},\mathbf{u}) + \nabla\varphi(\mathbf{x})^\top\mathbf{b} + \left\langle \nabla\varphi(\mathbf{x})\mathbf{u}^\top, \mathbf{A}\right\rangle + \sum_{i=1}^N \left(\left\langle \boldsymbol{\lambda}_i\mathbf{u}_i^\top, \mathbf{A}\right\rangle + \boldsymbol{\lambda}_i^\top\mathbf{b} - \boldsymbol{\lambda}_i^\top\widetilde{\mathbf{F}}_i + \nu_i\epsilon_i(\mathbf{x})\right) \\ &= \left\langle \nabla\varphi(\mathbf{x})\mathbf{u}^\top + \sum_{i=1}^N \boldsymbol{\lambda}_i\boldsymbol{u}_i^\top, \mathbf{A}\right\rangle + \left(\nabla\varphi(\mathbf{x}) + \sum_{i=1}^N \boldsymbol{\lambda}_i\right)^\top \mathbf{b} + \widehat{\dot{\varphi}}(\mathbf{x},\mathbf{u}) - \sum_{i=1}^N \left(\boldsymbol{\lambda}_i^\top\widetilde{\mathbf{F}}_i - \nu_i\epsilon_i(\mathbf{x})\right), \end{aligned}$$

where $(\boldsymbol{\lambda}_i, \nu_i) \in \mathcal{Q}_n$ are Lagrange multipliers corresponding to the constraints imposed by the data point $(\mathbf{x}_i, \mathbf{u}_i)$ and $\langle\cdot,\cdot\rangle$ denotes the trace inner product. The dual optimization problem is therefore:

$$\begin{aligned} \inf_{\substack{\boldsymbol{\lambda}_i\in\mathbb{R}^n\\ \nu_i\in\mathbb{R}}} \ & \widehat{\dot{\varphi}}(\mathbf{x},\mathbf{u}) - \sum_{i=1}^N \left(\boldsymbol{\lambda}_i^\top\widetilde{\mathbf{F}}_i - \nu_i\epsilon_i(\mathbf{x})\right) \\ \text{s.t.} \ & \sum_{i=1}^N \boldsymbol{\lambda}_i\mathbf{u}_i^\top = -\nabla\varphi(\mathbf{x})\mathbf{u}^\top \\ & \sum_{i=1}^N \boldsymbol{\lambda}_i = -\nabla\varphi(\mathbf{x}) \\ & \|\boldsymbol{\lambda}_i\|_2 \leq \nu_i \text{ for all } i \in 1,\ldots,N. \end{aligned}$$

For any $(\mathbf{x}_i, \mathbf{u}_i)$, choosing Lagrange multipliers such that $\nu_i > \|\boldsymbol{\lambda}_i\|_2$ increases the corresponding value of the problem compared to choosing $\nu_i = \|\boldsymbol{\lambda}_i\|_2$. We therefore simplify the problem as:

$$\begin{aligned}\inf_{\boldsymbol{\lambda}_i \in \mathbb{R}^n} \ & \widehat{\varphi}(\mathbf{x}, \mathbf{u}) - \sum_{i=1} \left(\boldsymbol{\lambda}_i^\top \widetilde{\mathbf{F}}_i - \|\boldsymbol{\lambda}_i\|_2 \epsilon_i(\mathbf{x}) \right) \\ \text{s.t.} \ & \sum_{i=1}^{N} \boldsymbol{\lambda}_i \mathbf{u}_i^\top = -\nabla\varphi(\mathbf{x})\mathbf{u}^\top \\ & \sum_{i=1}^{N} \boldsymbol{\lambda}_i = -\nabla\varphi(\mathbf{x}).\end{aligned}$$

This optimization problem can then be substituted in for the initial robust constraint in (DR-CCF-OP) to yield the final optimization problem (DR-CCF-SOCP). ■

Appendix B. Feasibility Proofs

In this appendix we provide proofs of the lemma and theorem presented on feasibility of the robust optimization problem in Section 4.

B.1. Proof of Lemma 1

Proof Consider the set $\mathcal{S}$, and define:

$$\mathcal{S}_{\mathbf{u}} \triangleq \left\{ \mathbf{u}_i \in \mathbb{R}^m \;\middle|\; \begin{bmatrix} \mathbf{u}_i \\ 1 \end{bmatrix} \in \mathcal{S} \right\}. \tag{19}$$

For arbitrary $\mathbf{x} \in \mathbb{R}^n$, let $(\mathbf{A}, \mathbf{b}) \in \mathcal{U}(\mathbf{x})$. By definition of the individual uncertainty sets $\mathcal{U}_i(\mathbf{x})$ and $\mathcal{U}(\mathbf{x})$ given in (14) and (15), respectively, we have that:

$$\left\| \mathbf{A}\mathbf{u}_i + \mathbf{b} - \widetilde{\mathbf{F}}_i \right\|_2 \leq \epsilon_i(\mathbf{x}), \tag{20}$$

for $\mathbf{u}_i \in \mathcal{S}_{\mathbf{u}}$. Taking the square of both sides and summing over the inputs in $\mathcal{S}_{\mathbf{u}}$ we arrive at:

$$\sum_{i=1}^{m+1} \left\| \mathbf{A}\mathbf{u}_i + \mathbf{b} - \widetilde{\mathbf{F}}_i \right\|_2^2 \leq \sum_{i=1}^{m+1} \epsilon_i(\mathbf{x})^2. \tag{21}$$

Defining $\mathbf{v}_i \triangleq \mathbf{A}\mathbf{u}_i + \mathbf{b} - \widetilde{\mathbf{F}}_i$, we can rewrite the left-hand side as:

$$\sum_{i=1}^{m+1} \left\| \mathbf{A}\mathbf{u}_i + \mathbf{b} - \widetilde{\mathbf{F}}_i \right\|_2^2 = \sum_{i=1}^{m+1} \sum_{j=1}^{n} |V_{ji}|^2 = \|\mathbf{V}\|_F^2, \tag{22}$$

where $\|\cdot\|_F$ is the Frobenius norm and $\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_{m+1} \end{bmatrix}$. Noting that $\|\mathbf{V}\|_F = \|\mathbf{V}^\top\|_F$, we have that:

$$\sum_{i=1}^{m+1} \left\| \mathbf{A}\mathbf{u}_i + \mathbf{b} - \widetilde{\mathbf{F}}_i \right\|_2^2 = \left\| \begin{bmatrix} \mathbf{v}_1^\top \\ \vdots \\ \mathbf{v}_{m+1}^\top \end{bmatrix} \right\|_F^2 = \left\| \begin{bmatrix} \mathbf{u}_1^\top \mathbf{A}^\top + \mathbf{b}^\top \\ \vdots \\ \mathbf{u}_{m+1}^\top \mathbf{A}^\top + \mathbf{b}^\top \end{bmatrix} - \begin{bmatrix} \widetilde{\mathbf{F}}_1^\top \\ \vdots \\ \widetilde{\mathbf{F}}_{m+1}^\top \end{bmatrix} \right\|_F^2. \tag{23}$$

By factoring the preceding term we arrive at:

$$\sum_{i=1}^{m+1} \left\| \mathbf{A}\mathbf{u}_i + \mathbf{b} - \widetilde{\mathbf{F}}_i \right\|_2^2 = \left\| \begin{bmatrix} \mathbf{u}_1^\top & 1 \\ \vdots & \vdots \\ \mathbf{u}_{m+1}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A}^\top \\ \mathbf{b}^\top \end{bmatrix} - \begin{bmatrix} \widetilde{\mathbf{F}}_1^\top \\ \vdots \\ \widetilde{\mathbf{F}}_{m+1}^\top \end{bmatrix} \right\|_F^2 \leq \sum_{i=1}^{m+1} \epsilon_i(\mathbf{x})^2. \tag{24}$$

Using the fact that $\|\mathbf{P}\|_2 \leq \|\mathbf{P}\|_F$ for any $\mathbf{P} \in \mathbb{R}^{m+1 \times n}$, we have that:

$$\left\| \begin{bmatrix} \mathbf{u}_1^\top & 1 \\ \vdots & \vdots \\ \mathbf{u}_{m+1}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A}^\top \\ \mathbf{b}^\top \end{bmatrix} - \begin{bmatrix} \widetilde{\mathbf{F}}_1^\top \\ \vdots \\ \widetilde{\mathbf{F}}_{m+1}^\top \end{bmatrix} \right\|_2^2 \leq \sum_{i=1}^{m+1} \epsilon_i(\mathbf{x})^2. \tag{25}$$

Taking the square root of both sides and employing the reverse triangle inequality we arrive at:

$$\left\| \begin{bmatrix} \mathbf{u}_1^\top & 1 \\ \vdots & \vdots \\ \mathbf{u}_{m+1}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A}^\top \\ \mathbf{b}^\top \end{bmatrix} \right\|_2 \leq \sqrt{\sum_{i=1}^{m+1} \epsilon_i(\mathbf{x})^2} + \left\| \begin{bmatrix} \widetilde{\mathbf{F}}_1^\top \\ \vdots \\ \widetilde{\mathbf{F}}_{m+1}^\top \end{bmatrix} \right\|_2. \tag{26}$$

Defining the matrix:

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}_1^\top & 1 \\ \vdots & \vdots \\ \mathbf{u}_{m+1}^\top & 1 \end{bmatrix}. \tag{27}$$

and noting that $\mathbf{U}$ has full rank by the linear independence of the vectors in the set $\mathcal{S}$, we have that:

$$\sigma_{m+1}(\mathbf{U}) \left\| \begin{bmatrix} \mathbf{A}^\top \\ \mathbf{b}^\top \end{bmatrix} \right\|_2 \leq \left\| \begin{bmatrix} \mathbf{u}_1^\top & 1 \\ \vdots & \vdots \\ \mathbf{u}_{m+1}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A}^\top \\ \mathbf{b}^\top \end{bmatrix} \right\|_2, \tag{28}$$

where $\sigma_{m+1}(\mathbf{U}) > 0$ is the smallest singular value of $\mathbf{U}$. Thus we may conclude that:

$$\left\| \begin{bmatrix} \mathbf{A}^\top \\ \mathbf{b}^\top \end{bmatrix} \right\|_2 \leq \frac{1}{\sigma_{m+1}(\mathbf{U})} \left(\sqrt{\sum_{i=1}^{m+1} \epsilon_i(\mathbf{x})^2} + \left\| \begin{bmatrix} \widetilde{\mathbf{F}}_1^\top \\ \vdots \\ \widetilde{\mathbf{F}}_{m+1}^\top \end{bmatrix} \right\|_2 \right), \tag{29}$$

proving that the uncertainty set $\mathcal{U}(\mathbf{x})$ is bounded. ■

B.2. Proof of Theorem 2

Proof The proof of Theorem 2 proceeds from the original structure of the robust control problem (DR-CCF-OP). The constraint that must be met in this controller by an input $\mathbf{u}^* \in \mathbb{R}^m$ is given by:

$$L_{\widehat{\mathbf{f}}}\varphi(\mathbf{x}) + \nabla\varphi(\mathbf{x})^\top \mathbf{b} + (L_{\widehat{\mathbf{g}}}\varphi(\mathbf{x}) + \nabla\varphi(\mathbf{x})^\top \mathbf{A})\mathbf{u}^* \leq -\alpha(\varphi(\mathbf{x})). \tag{30}$$

for all $(\mathbf{A}, \mathbf{b}) \in \mathcal{U}(\mathbf{x})$. Given the definitions of $\widetilde{\mathcal{U}}_\varphi(\mathbf{x})$ and $\mathcal{U}_\varphi(\mathbf{x})$ in (16) and (17), this constraint can equivalently be written as:

$$q + \mathbf{p}^\top \mathbf{u}^* \leq -\alpha(\varphi(\mathbf{x})) \tag{31}$$

for all $(\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x})$.

B.2.1. ONLY IF DIRECTION

To see the only if direction, assume that:

$$\mathcal{U}_\varphi(\mathbf{x}) \cap \mathcal{R} \neq \emptyset. \tag{32}$$

This implies that there exists a $(\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x})$ such that $q > -\alpha(\varphi(\mathbf{x}))$ and $\mathbf{p} = \mathbf{0}_m$. Thus for any choice of an input $\mathbf{u} \in \mathbb{R}^m$, we have that:

$$q + \mathbf{p}^\top \mathbf{u} = q > -\alpha(\varphi(\mathbf{x})), \tag{33}$$

violating the constraint in (31). Thus the optimization problem is infeasible.

B.2.2. IF DIRECTION

The other direction of the proof takes considerably more effort. To illuminate the various elements in the proof, we will proceed by proving a sequence of three statements of increasing strength: first when actuation is unnecessary for certification, then when there is an infinite gain margin, and finally in general.

Actuation unnecessary. For the first statement, we make the following assumption:

$$\forall\, (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}), \qquad \left\langle \begin{bmatrix} \mathbf{p} \\ q \end{bmatrix}, \begin{bmatrix} \mathbf{0}_m \\ 1 \end{bmatrix} \right\rangle \leq -\alpha(\varphi(\mathbf{x})), \tag{34}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product. This implies that for all $(\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x})$ we have $q \leq -\alpha(\varphi(\mathbf{x}))$. This implies that (31) is met for all $(\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x})$ with $\mathbf{u}^* = \mathbf{0}$. Intuitively, this assumption implies that for all possible Lie derivative pairs in $\mathcal{U}_\varphi(\mathbf{x})$, the certificate function condition is met with no input. Geometrically, this can be interpreted as the set $\mathcal{U}_\varphi(\mathbf{x})$ lying below the hyperplane $\mathcal{H} \subset \mathbb{R}^{m+1}$ defined by normal $\mathbf{n} = \begin{bmatrix} \mathbf{0}_m^\top & 1 \end{bmatrix}^\top$ and offset $-\alpha(\varphi(\mathbf{x}))$ as seen in Figure 4.

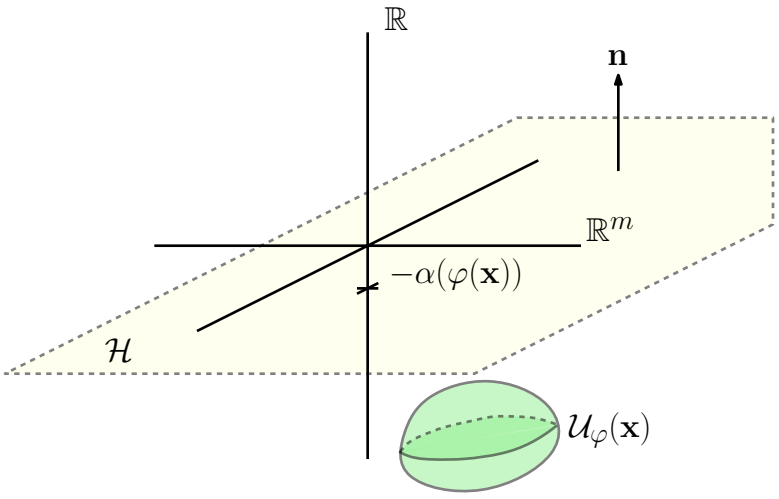

Figure 4. Geometric interpretation of the first statement in proving stability. The set $\mathcal{U}_\varphi(\mathbf{x})$ lies entirely below the hyperplane $\mathcal{H}$, implying the certificate function condition is met for all possible values of Lie derivatives with $\mathbf{u} = \mathbf{0}$.

Infinite gain margin. For the second statement, we begin by defining the projection function $\mathbf{\Pi} : \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^m$ such that for $(\mathbf{p}, q) \in \mathbb{R}^m \times \mathbb{R}$ we have:

$$\mathbf{\Pi}((\mathbf{p}, q)) = \mathbf{p}. \tag{35}$$

As $\mathbf{\Pi}$ is a linear transform, the image of $\mathcal{U}_\varphi(\mathbf{x})$ under $\mathbf{\Pi}$, defined as $\mathbf{\Pi}(\mathcal{U}_\varphi(\mathbf{x}))$, is compact and convex. We make the following assumption to prove the second statement:

$$\mathbf{\Pi}(\mathcal{U}_\varphi(\mathbf{x})) \cap \{\mathbf{0}_m\} = \emptyset. \tag{36}$$

Compared to the first statement, this allows for possible Lie derivative pairs $(\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x})$ with $q > -\alpha(\varphi(\mathbf{x}))$, such that $\mathbf{u} = \mathbf{0}$ is no longer a feasible solution. This assumption does imply that for all possible Lie derivative pairs $(\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x})$, $\mathbf{p} \neq \mathbf{0}$. Geometrically this appears in Figure 5, where now a portion of $\mathcal{U}_\varphi(\mathbf{x})$ is strictly above the hyperplane $\mathcal{H}$, but does not intersect the vertical axis. Given this assumption, and noting that both $\{\mathbf{0}_m\}$ and $\mathbf{\Pi}(\mathcal{U}_\varphi(\mathbf{x}))$ are compact and convex, we

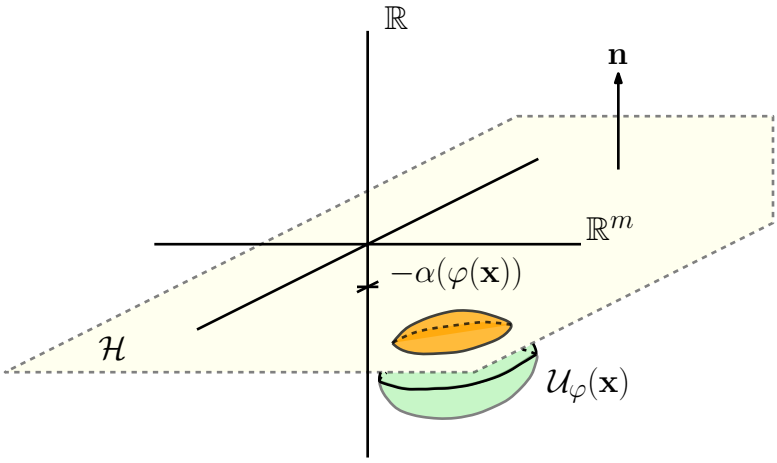

Figure 5. Geometric interpretation of the second statement in proving stability. The set $\mathcal{U}_\varphi(\mathbf{x})$ no longer lies entirely below the hyperplane $\mathcal{H}$, implying that input will be need to ensure the certificate function condition is met for all possible values of Lie derivatives.

may use the strict separating hyperplane theorem (Boyd and Vandenberghe, 2004) to conclude the existence of a hyperplane, $\mathcal{H}_\beta \subset \mathbb{R}^m$, defined by unit normal $\mathbf{s} \in \mathbb{R}^m$ and offset $\beta > 0$ such that:

$$\forall\, \mathbf{p} \in \mathbf{\Pi}(\mathcal{U}_\varphi(\mathbf{x})), \qquad \langle \mathbf{s}, \mathbf{p} \rangle \geq \beta. \tag{37}$$

As $\beta > 0$, a similar hyperplane $\mathcal{H}_0 \subset \mathbb{R}^m$ with the same normal $\mathbf{s}$ that passes through the origin can be defined. This can be further visualized by looking down the vertical axis, as seen in Figure 6, where the hyperplane $\mathcal{H}$ has been omitted.

These hyperplanes can be extended back into the ambient space $\mathbb{R}^m \times \mathbb{R}$ by defining hyperplanes $\mathcal{H}'_\beta \subset \mathbb{R}^m \times \mathbb{R}$ and $\mathcal{H}'_0 \subset \mathbb{R}^m \times \mathbb{R}$ with normal vector $\mathbf{s}' = \begin{bmatrix}\mathbf{s}^\top & 0\end{bmatrix}^\top$ and respective offsets β and 0. These hyperplanes serve to separate the vertical axis $\mathbf{0}_m \times \mathbb{R}$ from the cylinder defined by $\mathbf{\Pi}(\mathcal{U}_\varphi(\mathbf{x})) \times \mathbb{R}$, which contains $\mathcal{U}_\varphi(\mathbf{x})$. By definition we have that:

$$\forall\, (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}), \qquad \left\langle \mathbf{s}', \begin{bmatrix}\mathbf{p} \\ q\end{bmatrix} \right\rangle \geq \beta. \tag{38}$$

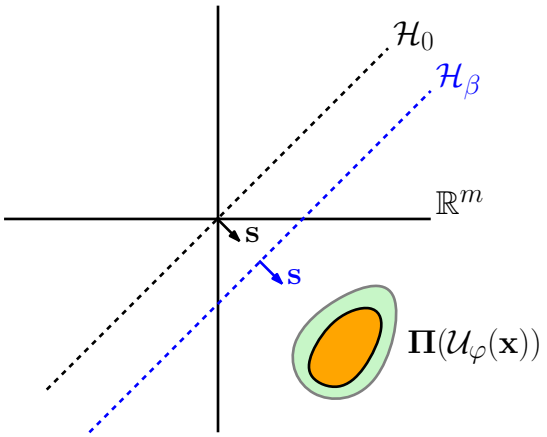

Figure 6. Top down view of the the second statement in proving feasibility. The green region corresponds to $\mathbf{\Pi}(\mathcal{U}_\varphi(\mathbf{x}))$ while the orange region highlights the portion of the set $\mathcal{U}_\varphi(\mathbf{x})$ lying above the hyperplane with normal $\mathbf{n}$ and offset $-\alpha(\varphi(\mathbf{x}))$. The line $\mathcal{H}_\beta$ represents the strictly separating hyperplane with normal $\mathbf{s}$ and offset β, while $\mathcal{H}_0$ is the hyperplane shifted to pass through the origin.

The function $\langle \mathbf{s}', \cdot \rangle : \mathcal{U}_\varphi(\mathbf{x}) \to \mathbb{R}$ is continuous on a compact domain, implying it obtains a minimum at a point $(\mathbf{p}^\star, q^\star) \in \mathcal{U}_\varphi(\mathbf{x})$ such that:

$$\left\langle \mathbf{s}', \begin{bmatrix} \mathbf{p}^\star \\ q^\star \end{bmatrix} \right\rangle \geq \beta > 0. \tag{39}$$

Similarly, the function $\left\langle \begin{bmatrix} \mathbf{0}_m^\top & 1 \end{bmatrix}^\top, \cdot \right\rangle : \mathcal{U}_\varphi(\mathbf{x}) \to \mathbb{R}$ is continuous on a compact domain, implying it obtains a maximum at a point $(\mathbf{p}^{\star\star}, q^{\star\star}) \in \mathcal{U}_\varphi(\mathbf{x})$ such that:

$$\forall\, (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}), \qquad \left\langle \begin{bmatrix} \mathbf{0}_m \\ 1 \end{bmatrix}, \begin{bmatrix} \mathbf{p}^{\star\star} \\ q^{\star\star} \end{bmatrix} \right\rangle = q^{\star\star} \geq q = \left\langle \begin{bmatrix} \mathbf{0}_m \\ 1 \end{bmatrix}, \begin{bmatrix} \mathbf{p} \\ q \end{bmatrix} \right\rangle. \tag{40}$$

We assume that $q^{\star\star} > -\alpha(\varphi(\mathbf{x}))$, or this setting reverts to the case proven in the first statement. Given these facts, our goal is to find an input $\mathbf{u}^* \in \mathbb{R}^m$ such that:

$$\forall\, (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}), \qquad q + \mathbf{p}^\top \mathbf{u}^\star \leq -\alpha(\varphi(\mathbf{x})), \tag{41}$$

implying $\mathbf{u}^*$ is a feasible solution for the robust optimization problem. To find an explicit value of $\mathbf{u}^*$ that satisfies this property, let us set $\mathbf{u}^* = -\gamma \mathbf{s}$ with $\gamma > 0$. Using the minimum found above, we have that:

$$\forall\, (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}), \qquad \langle \mathbf{s}, \mathbf{p} \rangle \geq \langle \mathbf{s}, \mathbf{p}^\star \rangle \geq \beta > 0 \tag{42}$$

or

$$\forall\, (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}), \qquad \langle -\gamma \mathbf{s}, \mathbf{p} \rangle \leq \langle -\gamma \mathbf{s}, \mathbf{p}^\star \rangle \leq -\gamma\beta < 0 \tag{43}$$

Using the maximum found above, we then have that:

$$\forall\, (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}), \qquad q + \mathbf{p}^\top \mathbf{u}^* = q - \gamma \mathbf{p}^\top \mathbf{s} \leq q^{\star\star} - \gamma\beta \tag{44}$$

Therefore if γ is chosen to satisfy:

$$\gamma \geq \frac{q^{\star\star} + \alpha(\varphi(\mathbf{x}))}{\beta}, \tag{45}$$

noting $\beta > 0$ by strict separation, we have the desired condition in (31), implying feasibility.

Another way to interpret this preceding result that is useful in proving the last statement begins with defining the set:

$$\mathcal{U}'_\varphi(\mathbf{x}, \mathbf{u}) = \left\{ (\mathbf{p}', q') \in \mathbb{R}^m \times \mathbb{R} \;\middle|\; \exists\, (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}) \text{ s.t. } (\mathbf{p}', q') = (\mathbf{p}, q + \mathbf{p}^\top \mathbf{u}) \right\}. \tag{46}$$

This set can be interpreted as a transformation of the set $\mathcal{U}_\varphi(\mathbf{x})$ that depends on an input $\mathbf{u}$, noting that $\mathcal{U}'_\varphi(\mathbf{x}, \mathbf{0}) = \mathcal{U}_\varphi(\mathbf{x})$. Further note that this is a linear transformation of the set $\mathcal{U}_\varphi(\mathbf{x})$ as we have:

$$\begin{bmatrix} \mathbf{p}' \\ q' \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{m \times m} & \mathbf{0}_m \\ \mathbf{u}^{\star\top} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ q \end{bmatrix}, \tag{47}$$

implying that $\mathcal{U}'_\varphi(\mathbf{x}, \mathbf{u})$ is compact and convex. If for some $\mathbf{u}^*$ the transformed set $\mathcal{U}'_\varphi(\mathbf{x}, \mathbf{u}^*)$ satisfies:

$$\forall\, (\mathbf{p}', q') \in \mathcal{U}'_\varphi(\mathbf{x}, \mathbf{u}^\star), \qquad \left\langle \begin{bmatrix} \mathbf{0}_m \\ 1 \end{bmatrix}, \begin{bmatrix} \mathbf{p}' \\ q' \end{bmatrix} \right\rangle \leq -\alpha(\varphi(\mathbf{x})), \tag{48}$$

such that it lies below the hyperplane $\mathcal{H}$, the input $\mathbf{u}^\star$ is a feasible solution for the robust optimization problem as:

$$\forall\, (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}), \qquad q + \mathbf{p}^\top \mathbf{u}^\star \leq -\alpha(\varphi(\mathbf{x})). \tag{49}$$

General case. The third statement requires only the main assumption in Theorem 2 such that:

$$\mathcal{U}_\varphi(\mathbf{x}) \cap \mathcal{R} = \emptyset. \tag{50}$$

In particular, we no longer require that:

$$\mathbf{\Pi}(\mathcal{U}_\varphi(\mathbf{x})) \cap \{\mathbf{0}_m\} = \emptyset. \tag{51}$$

To proceed, we define the following set:

$$\overline{\mathcal{U}}_\varphi(\mathbf{x}) = \left\{ (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}) \;\middle|\; \left\langle \begin{bmatrix} \mathbf{0}_m \\ 1 \end{bmatrix}, \begin{bmatrix} \mathbf{p} \\ q \end{bmatrix} \right\rangle \geq -\alpha(\varphi(\mathbf{x})) \right\}, \tag{52}$$

which corresponds to the Lie derivative pairs in the set $\mathcal{U}_\varphi(\mathbf{x})$ that do not meet the certificate function condition under no input. As $\overline{\mathcal{U}}_\varphi(\mathbf{x})$ is defined as an intersection of the convex set $\mathcal{U}_\varphi(\mathbf{x})$ and a half-space, it is also convex. We consider the projection of this set, which is convex and by assumption satisfies:

$$\mathbf{\Pi}(\overline{\mathcal{U}}_\varphi(\mathbf{x})) \cap \{\mathbf{0}_m\} = \emptyset. \tag{53}$$

Once again we may use the strict separating hyperplane theorem, this time to separate the set $\mathbf{\Pi}(\overline{\mathcal{U}}_\varphi(\mathbf{x}))$ from $\{\mathbf{0}_m\}$ with the hyperplane $\mathcal{H}_\beta$ with unit normal $\mathbf{s}$ and offset β. This hyperplane can also be shifted to pass through the origin, given by $\mathcal{H}_0$. This results in the configuration seen in Figure 7. These hyperplanes can be extended back into the ambient space $\mathbb{R}^m \times \mathbb{R}$ by defining hyperplanes $\mathcal{H}'_\beta \subset \mathbb{R}^m \times \mathbb{R}$ and $\mathcal{H}'_0 \subset \mathbb{R}^m \times \mathbb{R}$ with normal vector $\mathbf{s}' = \begin{bmatrix} \mathbf{s}^\top & 0 \end{bmatrix}^\top$ and respective offsets β and 0. These hyperplanes serve to separate the vertical axis $\mathbf{0}_m \times \mathbb{R}$ from the cylinder defined by $\mathbf{\Pi}(\overline{\mathcal{U}}_\varphi(\mathbf{x})) \times \mathbb{R}$. By definition we have that:

$$\forall\, (\mathbf{p}, q) \in \overline{\mathcal{U}}_\varphi(\mathbf{x}), \qquad \left\langle \mathbf{s}', \begin{bmatrix} \mathbf{p} \\ q \end{bmatrix} \right\rangle \geq \beta. \tag{54}$$

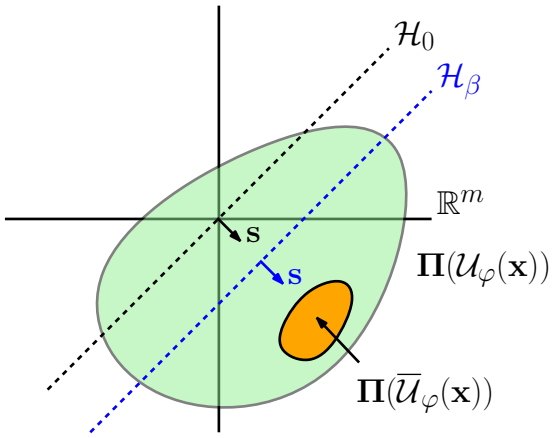

Figure 7. Top down view of the third statement in proving feasibility. The green region corresponds to $\mathbf{\Pi}(\mathcal{U}_\varphi(\mathbf{x}))$ while the orange region highlights the set $\overline{\mathcal{U}}_\varphi(\mathbf{x})$ lying above the hyperplane $\mathcal{H}$. The line $\mathcal{H}_\beta$ represents the strictly separating hyperplane with normal $\mathbf{s}$ and offset β, while $\mathcal{H}_0$ is the hyperplane shifted to pass through the origin.

The function $\langle \mathbf{s}', \cdot \rangle : \overline{\mathcal{U}}_\varphi(\mathbf{x}) \to \mathbb{R}$ is continuous on a compact domain, implying it obtains a minimum at a point $(\mathbf{p}^\star, q^\star) \in \overline{\mathcal{U}}_\varphi(\mathbf{x})$ such that:

$$\left\langle \mathbf{s}', \begin{bmatrix} \mathbf{p}^\star \\ q^\star \end{bmatrix} \right\rangle \geq \beta > 0. \tag{55}$$

Similarly, the function $\left\langle \begin{bmatrix} \mathbf{0}_m^\top & 1 \end{bmatrix}^\top, \cdot \right\rangle : \overline{\mathcal{U}}_\varphi(\mathbf{x}) \to \mathbb{R}$ is continuous on a compact domain, implying it obtains a maximum at a point $(\mathbf{p}^{\star\star}, q^{\star\star}) \in \overline{\mathcal{U}}_\varphi(\mathbf{x})$ such that:

$$\forall\, (\mathbf{p}, q) \in \overline{\mathcal{U}}_\varphi(\mathbf{x}), \qquad \left\langle \begin{bmatrix} \mathbf{0}_m \\ 1 \end{bmatrix}, \begin{bmatrix} \mathbf{p}^{\star\star} \\ q^{\star\star} \end{bmatrix} \right\rangle = q^{\star\star} \geq q = \left\langle \begin{bmatrix} \mathbf{0}_m \\ 1 \end{bmatrix}, \begin{bmatrix} \mathbf{p} \\ q \end{bmatrix} \right\rangle. \tag{56}$$

We assume that $q^{\star\star} > -\alpha(\varphi(\mathbf{x}))$, or this setting reverts to the case proven in the first statement. As before we may choose an input $\mathbf{u}^* = -\gamma \mathbf{s}$ with:

$$\gamma \geq \frac{q^{\star\star} + \alpha(\varphi(\mathbf{x}))}{\beta}. \tag{57}$$

Unlike before, this only ensures that the condition (31) is met for $(\mathbf{p}, q) \in \overline{\mathcal{U}}_\varphi(\mathbf{x})$, rather than $(\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x})$. To see that this presents an additional challenge, let us define the two following sets:

$$\mathcal{H}_0^+ \triangleq \left\{ (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}) \;\middle|\; \left\langle \mathbf{s}', \begin{bmatrix} \mathbf{p} \\ q \end{bmatrix} \right\rangle > 0 \right\} \tag{58}$$

$$\mathcal{H}_0^- \triangleq \left\{ (\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}) \;\middle|\; \left\langle \mathbf{s}', \begin{bmatrix} \mathbf{p} \\ q \end{bmatrix} \right\rangle < 0 \right\} \tag{59}$$

For a point $(\mathbf{p}, q) \in \mathcal{H}_0^-$ we have that:

$$\mathbf{p}^\top \mathbf{u}^* = -\gamma \mathbf{p}^\top \mathbf{s} > 0, \tag{60}$$

or that the input $\mathbf{u}^*$ is working against meeting the certificate function condition for $(\mathbf{p}, q)$. The hyperplane $\mathcal{H}'_0$ separates points for which $\mathbf{u}^*$ improves or worsens the certificate condition, as illustrated in Figure 8. We must then verify that this choice of $\mathbf{u}^*$ does not lead to violations of the condition (31) for $(\mathbf{p}, q) \notin \overline{\mathcal{U}}_\varphi(\mathbf{x})$.

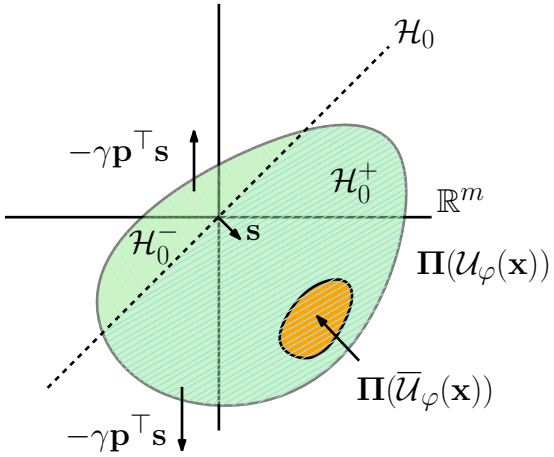

Figure 8. Top down view of the the third statement in proving feasibility. The blue rising hatched region corresponds to $\mathcal{H}_0^+$ for which the input $\mathbf{u}^* = -\gamma \mathbf{s}$ works to improve the certificate condition. The pink falling hatched region corresponds to $\mathcal{H}_0^-$ for which the input $\mathbf{u}^* = -\gamma \mathbf{s}$ worsens the certificate condition. They are separated by the hyperplane $\mathcal{H}_0$.

To this end we define the set:

$$\overline{\mathcal{U}}'_\varphi(\mathbf{x}, \mathbf{u}) = \left\{ (\mathbf{p}', q') \in \mathcal{U}'_\varphi(\mathbf{x}, \mathbf{u}) \;\middle|\; \left\langle \begin{bmatrix} \mathbf{0}_m \\ 1 \end{bmatrix}, \begin{bmatrix} \mathbf{p}' \\ q' \end{bmatrix} \right\rangle \geq -\alpha(\varphi(\mathbf{x})) \right\}, \tag{61}$$

with $\overline{\mathcal{U}}'_\varphi(\mathbf{x}, \mathbf{0}) = \overline{\mathcal{U}}_\varphi(\mathbf{x})$. The set $\overline{\mathcal{U}}'_\varphi(\mathbf{x}, \mathbf{u})$ contains the points in $\mathcal{U}'_\varphi(\mathbf{x}, \mathbf{u})$ that are in or above the hyperplane $\mathcal{H}$. We note that it is a convex set, and that by strict separation, we have:

$$\overline{\mathcal{U}}'_\varphi(\mathbf{x}, \mathbf{0}) \cap \mathcal{H}'_0 = \emptyset. \tag{62}$$

Furthermore, for any value of γ, we have:

$$\overline{\mathcal{U}}'_\varphi(\mathbf{x}, -\gamma \mathbf{s}) \cap \mathcal{H}'_0 = \emptyset. \tag{63}$$

This implies that for any choice of γ, any point $(\mathbf{p}, q) \in \mathcal{U}_\varphi(\mathbf{x}) \cap \mathcal{H}'_0$ will satisfy condition (31). Likewise, considering a point $(\mathbf{p}, \mathbf{q}) \in \mathcal{H}_0^+ \setminus \overline{\mathcal{U}}'_\varphi(\mathbf{x}, \mathbf{0})$ (for which condition (31) is met with no input and the satisfaction of the certificate condition improves), for all $\gamma > 0$ we have that:

$$\mathbf{p}^\top \mathbf{u}^* < 0, \tag{64}$$

implying that condition (31) is met for $(\mathbf{p}, \mathbf{q})$. By continuously varying γ, we may conclude that there exists $\gamma^* \in \mathbb{R}$ satisfying:

$$0 < \gamma^* \leq \frac{q^{\star\star} + \alpha(\varphi(\mathbf{x}))}{\beta}, \tag{65}$$

such that:

$$\max\left\{q' \in \mathbb{R} \;\middle|\; (\mathbf{p}', q') \in \overline{\mathcal{U}}'_{\varphi}(\mathbf{x}, -\gamma^*\mathbf{s})\right\} = -\alpha(\varphi(\mathbf{x})). \tag{66}$$

Importantly, for $\gamma = \gamma^*$, we have that all points $(\mathbf{p}, q) \in \mathcal{H}_0^+$ satisfy condition (31), and that at least one point, denoted $\mathbf{v}^+ \in \overline{\mathcal{U}}'_{\varphi}(\mathbf{x}, -\gamma^*\mathbf{s}) \cap \mathcal{H}_0^+$, satisfies it with equality. Now assume for contradiction that:

$$\overline{\mathcal{U}}'_{\varphi}(\mathbf{x}, -\gamma^*\mathbf{s}) \cap \mathcal{H}_0^- \neq \emptyset, \tag{67}$$

or that there exists $\mathbf{v}^- \in \overline{\mathcal{U}}'_{\varphi}(\mathbf{x}, -\gamma^*\mathbf{s}) \cap \mathcal{H}_0^-$. As the set $\overline{\mathcal{U}}'_{\varphi}(\mathbf{x}, -\gamma^*\mathbf{s})$ is convex, we have that:

$$\forall\, \lambda \in [0, 1], \qquad \mathbf{v}(\lambda) \triangleq \lambda\mathbf{v}^+ + (1-\lambda)\mathbf{v}^- \in \overline{\mathcal{U}}'_{\varphi}(\mathbf{x}, -\gamma^*\mathbf{s}). \tag{68}$$

For any value of $\lambda \in [0, 1]$, taking the inner product of $\mathbf{v}(\lambda)$ with the normal vector $\mathbf{s}'$ defining $\mathcal{H}'_0$ yields:

$$\langle \mathbf{s}', \mathbf{v}(\lambda)\rangle = \lambda\langle \mathbf{s}', \mathbf{v}^+\rangle + (1-\lambda)\langle \mathbf{s}', \mathbf{v}^-\rangle. \tag{69}$$

Noting that $\langle \mathbf{s}', \mathbf{v}^+\rangle > 0$ and $\langle \mathbf{s}', \mathbf{v}^-\rangle < 0$, let us define:

$$\lambda^* = -\frac{\langle \mathbf{s}', \mathbf{v}^-\rangle}{\langle \mathbf{s}', \mathbf{v}^+\rangle - \langle \mathbf{s}', \mathbf{v}^-\rangle} \in (0, 1). \tag{70}$$

It follows that:

$$\langle \mathbf{s}', \mathbf{v}(\lambda^*)\rangle = 0, \tag{71}$$

implying $\mathbf{v}(\lambda^*) \in \overline{\mathcal{U}}'_{\varphi}(\mathbf{x}, -\gamma^*\mathbf{s}) \cap \mathcal{H}'_0$. This contradicts the fact:

$$\overline{\mathcal{U}}'_{\varphi}(\mathbf{x}, -\gamma^*\mathbf{s}) \cap \mathcal{H}'_0 = \emptyset, \tag{72}$$

implying:

$$\overline{\mathcal{U}}'_{\varphi}(\mathbf{x}, -\gamma^*\mathbf{s}) \cap \mathcal{H}_0^- = \emptyset, \tag{73}$$

or that all points $(\mathbf{p}, q) \in \mathcal{U}_{\varphi}(\mathbf{x})$ satisfy condition (31) for the input $\mathbf{u}^* = -\gamma^*\mathbf{s}$, ensuring feasibility. ■